UCRL-JC-134369
PREPRINT

Support and Development for Remote Collaborations in Fusion Research

T. A. Casper
R. A. Jong
W. H. Meyer
J. M. Moller

This paper was prepared for submittal to the
2nd International Atomic Energy Agency Technical Committee Meeting on
Control, Data Acquisition and Remote Participation for Fusion Research
Lisbon, Portugal
July 19-21, 1999

July 15, 1999

Support and Development for Remote Collaborations in Fusion Research

T.A. Casper, R.A. Jong, W.H. Meyer and J.M. Moller
Lawrence Livermore National Laboratory
7000 East Ave. P.O. Box 808 L637
Livermore, CA 94550
(925) 422-0787, fax (925) 424-6401, casper1@llnl.gov

Abstract

Major fusion experiments and modeling efforts rely on joint research of scientists from several locations around the world. A variety of software tools are in use to provide remote interactive access to facilities and data are routinely available over wide-area-network connections to researchers. Audio and video communications, monitoring of control room information and synchronization of remote sites with experimental operations all enhance particpation during experiments. Remote distributed computing capabilities allow utilization of off-site computers that now help support the demands of control room analyses and plasma modeling. A collaborative software development project is currently using object technologies with CORBA-based communications to build a network executable transport code that further demonstrates the ability to utilize geographically dispersed resources. Development to extend these concepts with security and naming services and possible applications to instrumentation systems has been initiated. An Information Technology Initiative is deploying communication systems, ISDN (telephone) and IP (network) audio/video (A/V) and web browser-based, to build the infrastructure needed to support remote physics meetings, seminars and interactive discussions.

1. Remote Experimentation Support

Collaboration on magnetic fusion (MFE) experiments worldwide has a long history of participation by multiple research groups, mostly by travel to experimental sites. Advances in desktop computing power and high performance networks now support real-time interactive research. Development of collaborative or virtual laboratories provides the infrastructure to support a more distributed work force. Design of our Remote Experimental Site concept [1-2] was a direct outgrowth and extension of networking applications used for local control and data acquisition systems on fusion and free electron laser experiments at LLNL [3]. Stimulated by involvement on international scale experiments, both the Alcator C-mod [4] and DIII-D [5] tokamaks were operated from LLNL (distances of ~5000kM and ~800kM respectively) and demonstrated the degree to which experimental control is possible with present day systems.

High-performance wide-area computational networks are necessary for utilizing distributed computing and file services, allow for near real-time synchronization of operations among geographically distant sites, and provide bandwidth for audio and video communications. A testbed virtual laboratory, developed under the Distributed Collaborative Experimental Environments (DCEE, [6-8]) Initiative, was realized at the DIII-D National Fusion Facility and is actively used by LLNL during DIII-D operations. The Distributed Computing Environment (DCE) provides a secure, wide-area-network (WAN) computing cluster with the Distributed File Services (DFS, derived from AFS) providing a shared, global file system. A library of services supports inter-process communications (IPCS, [9]) over the wide area, heterogeneous network upon that is built synchronization and process queue control mechanisms. We find that simplified software maintenance and use and mirroring of data can be achieved with DFS but its use has not yet achieved widespread acceptance at fusion sites. The main features of this virtual laboratory supporting LLNL site operations on DIII-D are:

- Distributed computations on several workstations at LLNL and General Atomics (GA)
- Multicast backbone (MBONE) audio/video monitoring of the control room operations
- Remote data access during operations and for post-experiment analysis and modeling
- Java-Chat interface for on-line discussions and a log of operations
- Web-based operations monitoring tools

During most experiments on DIII-D, equilibrium analysis is almost entirely provided by a distributed version of EFIT (http://lithos.gat.com/efit) optimized for data access and network performance and running on clients over the wide area network connection between LLNL and GA. Two interfaces to the distributed EFIT are available, a web [5-8] user interface for submitting post-experiment processing tasks and a new python-language graphical interface developed for use in the control room, Figure 1.

The DIII-D data library, PTDATA [10], used during acquisition is very efficient and provides much of the data access for intershot analysis tasks running on- and off-site as data are being acquired. IPCS messaging, now used for local and remote notification of data events and control of processing, provides fine-grain synchronizing of analysis tasks that helps minimize processing response time. Processed results, including those from remote sites, are sent to the new DIII-D version of the MDSplus (http://www.pfc.mit.edu/mdsplus) data system. This provides a more unified repository for analyzed data accessible by off-site participants. Use of the Interactive Display Language (IDL, [11]) has grown to the extent that it has become a standard scripting language for the development of data analysis tools. Since many of the data manipulations needed for analysis are quite similar among experiments, development of such tools are beginning to be used by some of the smaller experimental programs. For example, at the new LLNL Sustained Spheromak Experiment, SSPX, we utilize the ReviewPlus [12] data display tool developed at DIII-D, Figure 2.

Audio and video monitoring and remote meeting support has mostly been done via MBONE/sdr (http://www.mbone.com) connections that provide reasonably good, bi-directional A/V communications. However, it requires a moderate level of technical support to maintain, is not uniformly available at all sites and not allowed at all on some local networks. We are currently in the process of converting experimental monitoring to the commercially available RealNetworks (http://www.real.com) implementation, Figure 2. This should provide more ubiquitous connections among a greater variety and performance level of computers at the expense of bi-directional communications as it is a "transmit only" technology. Electronic notebooks are under development for experimental use but they currently lack the real time communication features of the Java/IRC-Chat (http://www.irchelp.org) interface we rely on.

To date, the development of secure techniques for interacting with data and instrumentation systems has not received sufficient attention. Existing controls and data systems generally do not provide for authorization, validation of access to categories of services. Nearly all analysis tools, even those that write results back into data files, lack security features other than those provided by entirely local use or by standard password protection that is not acceptable at many laboratories. Without the addition of such security features, it will become increasingly difficult to support collaborative environments at nuclear capable sites, such as those with DT-fusion experiments. We are now exploring increased use of secure architectures to either develop new interfaces or "wrap" existing applications. Secure shell (SSH) connections currently provide encrypted access to applications, data and controls. DCE applications can be run with the Kerberos-based security server. A DCE version of the distributed EFIT code was implemented as a first test of secure architectures for WAN-based analysis. We are exploring the use of object technologies based on CORBA-communications (Common Object Request Broker Architecture) to establish security and naming services for fusion applications. As collaborations continue to grow and become even more dispersed, authorization becomes increasingly important to prevent multiple remote users from simultaneously accessing a given resource, for example, control of an instrument or to protect the integrity and property rights of certain forms of data.

2. Distributed modeling code development

The National Transport Code Collaboration (NTCC) Demonstration Project [13] is a collaboration of physicists and computer scientists from national laboratories, universities and private industry. It's goal is to produce a library of transport code modules that meet a set of well-defined standards and to construct a demonstration code using modern computing techniques with a flexible framework that permits the rapid assembly of customized, web-invocable transport codes. The NTCC is a prime example of a collaboration between geographically dispersed scientists working together towards a common well-defined goal, development of a network-distributed application code to support collaborations. While NTCC development applies to transport modeling, we are interested in generalizing and extending it to data analysis and control system applications.

The demonstration code uses a mix of modules written in the Python, C++, C, and Java programming languages. A C++ data server provides network access to data with display capabilities and to run a Python physics server, Figure 3. The physics server executes a steerable application employing both new object modules and legacy transport modules written in Fortran 77. Clients access applications using a web-invocable, Java-based graphical user interface (GUI), while the network communications between the clients and servers are all based on CORBA and an appropriate interface definition language interface. The CORBA structure enables the data server, the physics server and GUI clients to reside on systems with diverse operating systems. GUI clients run on Linux, Solaris, HP, Windows 95 and NT, IRIX, and Macintosh operating systems while the data and physics servers run on Linux, IRIX, HP, and Solaris based platforms. All servers and clients run at widely separated geographic locations spanning the entire United States. The collaborative code development relies on the use of the Concurrent Version System software that enables widely dispersed developers to work independently and still maintain version control on the demonstration code as a whole.

Using the CORBA structure in an object oriented programming environment, the NTCC Demonstration Project has demonstrated desktop access, through a browser, to experimental data and physics analysis applications residing at remote sites. A flexible client provides data access and interactive control over the physics server, both with display capability. The physics server can modify parameters and interactively control the run, thus providing a steerable method for running the transport calculations. The data server currently provides data from multiple sources at a single site, but eventually will provide data from multiple sources at local and remote sites.

3. Information Technology Development

We have recently turned our attention to improving support for remote attendance at scientific seminars, program management and working-group meetings. Due mostly to a lack of infrastructure, the quality and capabilities for remote collaborative meeting support have been limited and, thus, the culture of remote collaboration has been slower to develop. Under the umbrella of the Virtual Laboratory for Technology, the Office of Fusion Energy Science began an Information Technology Initiative aimed at improving communication technologies for scientific collaborations in fusion research. A working group representing the major experiments and laboratories reviewed available technologies and provided information and recommendations concerning the nature and amount of infrastructure required to significantly improve the quality of collaborative meetings and seminars.

With the initial focus on the three main experiments in the US MFE program; DIII-D, Alcator C-mod and NSTX, plus the Virtual Laboratory for Technology (VLT) that includes some inertial fusion (IFE) needs, funding was provided to begin to build the infrastructure. Due to the great number of collaborating institutions currently working on these experiments, a significant portion of the fusion energy program is covered. A bulk purchase of network capable projection systems, the ShowStationIP (http://www.polycom.com) was completed and hardware distributed to 23 sites around the country, both laboratories and universities. In addition to locally projecting presentations, these systems support the T.120 standard for document conferencing that allows transmission and simultaneous viewing of presentations with integrated multi-point bridging for up to 15 sites. An integrated operating system and web server automatically sends the presentation to connected web browsers. Remote participants can upload presentations and show them using password protected connections. The advantage of this technology is that it provides high quality (jpeg image) viewing of presentations, sketches, etc., Figure 4, as compared to standard video without a loss of interactivity during the meeting.

We are now reviewing and upgrading video conferencing capabilities to further support remote collaborative meetings. Several groups are formulating plans to purchase videoconferencing systems for installation at many US fusion sites. Videoconferencing standards are sufficiently robust that little difficulty is experienced with ISDN interoperability. At the LLNL site, we have installed a system that supports both ISDN and IP videoconferencing (ViewStation, http://www.polycom.com) at up to full motion video (30 frames/sec) or multi-point (up to four) connections at reduced rates without connecting to an external videoconferencing bridge. It is capable of sending output to web-browsers and to RealServers for broadcasts to more isolated locations, Figure 4. With the combination of presentation material and video conferencing to the web, we will be able to support even single researcher and home connections to our meetings. We also envision using this videoconferencing hardware, operating in IP-mode with connections to RealServer technology, as a commercially supported, high-quality replacement for MBONE control room A/V monitoring. The videoconferencing provides the return path for any discussions with control room staff.

Web-browser applications are being developed to publicize and track the use of this new technology. The PPPL group maintains a web site, http://www.fusionscience.org, as a focal point for this purpose. The MIT group has assembled a prototype web page to publicize the collaboration capabilites at each program, http://www.fusionscience.org /cit/cit_fusion.html. At LLNL, we have put together a prototype URL for submitting and maintaining a schedule of meetings on a fusion-program-wide basis, Figure 5, that is temporarily located at http://skibuff.llnl.gov/DOE/DOE_Video_Schedule.html.

Even with the technology only recently installed and not fully operational at all sites, use is rapidly growing. Meetings are now routinely held in our remote collaboration conference room at LLNL, a small room with an early version of the ShowStation and both a ViewStation for ISDN and IP conferencing and a workstation supporting MBONE. The DIII-D group at GA broadcasts a weekly integrated modeling meeting that is routinely attended by MFE experimentalists and theorists at LLNL working on DIII-D. Similarly, the local IFE group, split between the Lawrence Livermore and Lawrence Berkeley laboratories, has their weekly program meetings via this conference room. From its very formation as a national collaborative development, the NTCC group has quickly embraced the collaboration technologies rather than make several trips a year to a common meeting location. This technical collaboration between scientists involves infrequent face-to-face meetings, but more frequent conference telephone calls and e-mail correspondence. Recently, with some new information technology barely in place, a multi-site conference that spanned three days was held. The conference, with a moderator and main connection based at Lehigh University in Bethlehem, PA, included remote participants from LLNL in Livermore, CA, GA in La Jolla, CA, The University of Texas (UT) in Austin, TX, and the Oak Ridge National Laboratory (ORNL) in Oak Ridge, TN. Visual presentations were made and simultaneously viewed from connections using ShowStations at Lehigh, LLNL, and GA and web browsers at Lehigh, GA, UT, and ORNL. At the time of this conference, video conferencing was not available at many sites, so telephone conference calls supported the audio. While these audio communications between widely separated participants were acceptable and sufficient for simple information transfer, being able to visually interact with a remote participant, even at a reduced frame rate, is better from an information transfer point of view and also from a cultural perspective.

4. Summary

Efforts at increasing the ease and flexibility for remote collaborations are ongoing in the fusion program. Between LLNL and the DIII-D National Fusion Facility, remote participation in experiments and data analysis are a daily routine. The LLNL site is fairly well integrated into the operations and IPCS messaging provides near real-time synchronization of monitoring and processing tasks. Access to data is rapid and robust and allows for production level data processing at LLNL in direct support of operations. Improvements in the A/V communications are planned as well as better access to data analysis code. Similar capabilities are available at Alcator C-mod and are being developed for the NSTX experimental program. While MBONE tools have worked reasonably well, we are heading toward commercial web-based systems for better access to universities and homes. The use of object technologies and CORBA communications have been demonstrated by the NTCC group to the degree that we are now pursuing their use for data analysis and instrumentation applications. While security remains rudimentary for data and controls systems, password protection and SSH, we are investigating secure architectures under both DCE and CORBA. The recent Information Technology Initiative is putting communications infrastructure hardware in the community and is already having a positive affect on our ability to hold joint meetings with remote participants.

5. Acknowledgements

We give special thanks to the DIII-D group for their support and willingness to deploy many of the information technologies in the experimental environment and to D. Schissel and B. McHarg and the DIII-D computations group in particular. We would like to acknowledge the efforts of the entire NTCC Demonstration Project under the direction of Dr. Arnold Kritz. NTCC, funded by the US DOE, is a collaboration of physicists and computer scientists from National Laboratories (LLNL*, ORNL, PPPL), universities (University of Colorado, Lehigh University, Massachusetts Institute of Technology, University of Texas)

*under contract No. W-7405-Eng-48

and private industry (General Atomics and Tech-X Corporation). Finally, we acknowledge the intensive efforts by several members of the special working group for the Information Technology Initiative: Walter Sadowski and Kyle Norovsky at the US DOE, Steve Sabbaugh (NSTX), Martin Greenwald (Alcator C-Mod), Dave Schissel (DIII-D) and Pat Stewart (VLT) for their efforts at getting information technologies deployed in the US fusion program.

References

[1] T.A. Casper, "Remote Experimental Site Concept for Diagnostic Collaborations in Fusion", International School of Plasma Physics, Workshop on Diagnostics for Contemporary Fusion Experiments, Varenna, Italy, September, 1991

[2] T. A. Casper and W. J. Lennon, "Remote Experimental Site: A Command and Analysis Center for Big Physics Experimentation", Proceedings of the 14th IEEE/NPSS Symposium on Fusion Engineering, 1991, p. 704.

[3] D.N. Butner, M. Drlik, W.H. Meyer, J.M. Moller and G.G. Preckshot, "An Integrated, Multi-vendor Distributed Data Acquisition System", Rev. Sci. Inst., **59** (1786) 1988.

[4] S. Horne et al., *"Remote Control of Alcator C-Mod from Lawrence Livermore National Laboratory"*, Fusion Technology, 32,152 (1997).

[5] T. A. Casper, et al., "Collaboratory Operations in Magnetic Fusion Scientific Research", Interactions, **3** (56) May/June 1998.

[6] T. A. Casper, et al. , "Remote Experimental Environment: Building a Collaboratory for Fusion Research", Computers in Physics, **12** (220), May/June 1998.

[7] T.A. Casper, W.H. Meyer and J.M. Moller, The Collaboratory for Support of Scientific Research, Proceedings of the Workshop on Remote Participation in Fusion Experiments, June 28-29, 1998, Prague, Czech Republic.

[8] B.B. McHarg Jr., T.A. Casper, S. Davis, D. Greenwood, Tools for remote collaboration on the DIII-D National Fusion Facility, Fusion Engineering Design, 43 (1999) 343-355.

[9] G.G. Preckshot and D.N. Butner, "A Simple and Efficient Interprocess Communications System for Actually Using a Laboratory Network", IEEE Trans. Nuclear Science **NS-34** (858) 1987.

[10] B. McHarg, Access to DIII-D Data Located in Multiple Files and Multiple Locations, Proceedings of 15th IEEE/NPSS Symposium. on Fusion Engineering, p. 123, 1993.

[11] IDL is a registered trademark of Research Systems, Inc., Boulder, CO.

[12] J. Schachter, Q. Peng and D.P. Schissel, "Data analysis software tools for enhanced collaboration at the DIII-D national fusion facility, presentation O12, this meeting.

[13] J.R. Cary et al., Client-Server for Remote Collaboration and Remote Invocation of Analysis Applications and Application to the National Transport Code, Proceedings of the Workshop on Remote Participation in Fusion Experiments, June 28-29, 1998, Prague, Czech Republic.

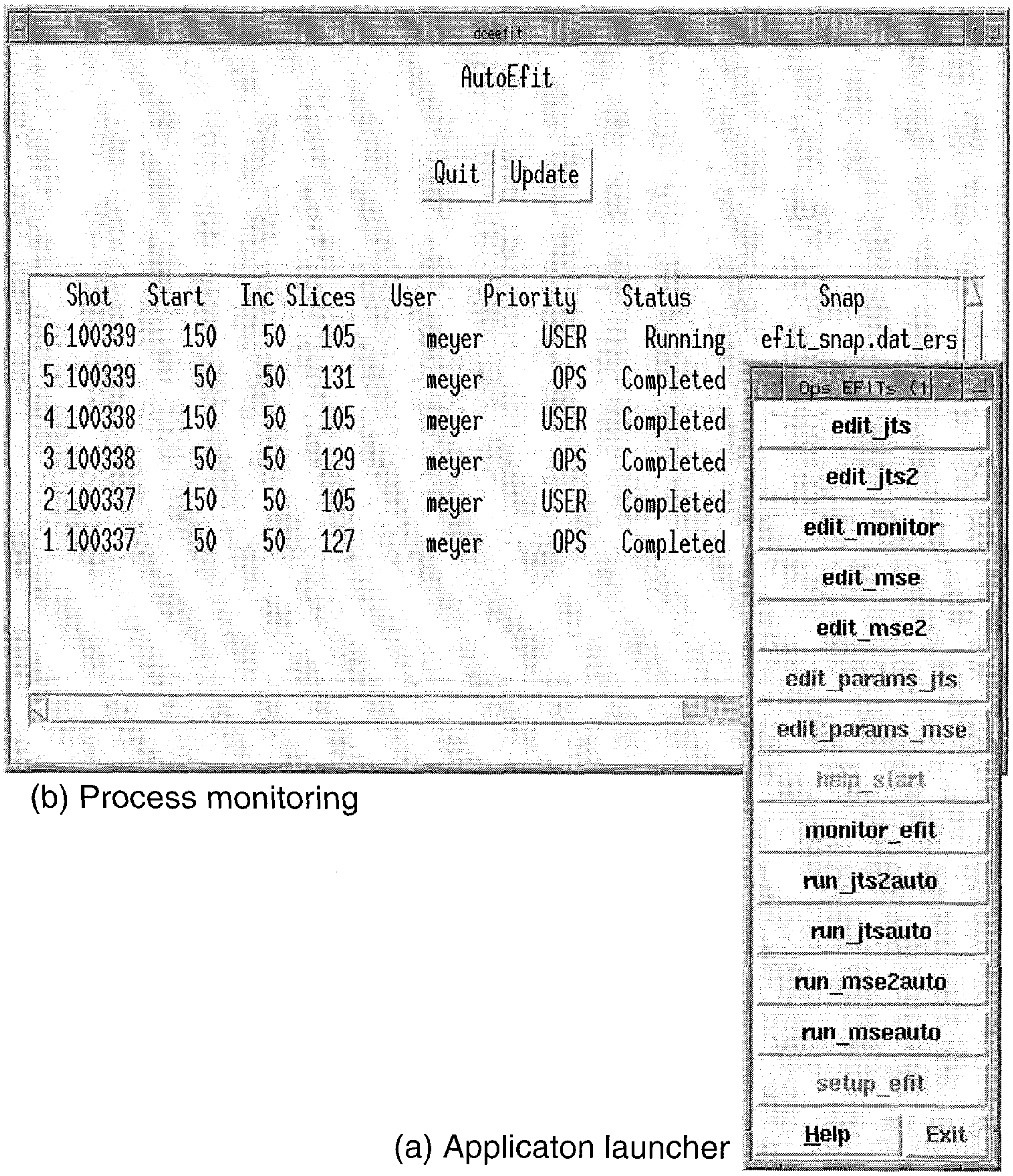

Figure 1 Python-language GUI interface for DCE EFIT analysis code: (a) application launcher and (b) process monitor.

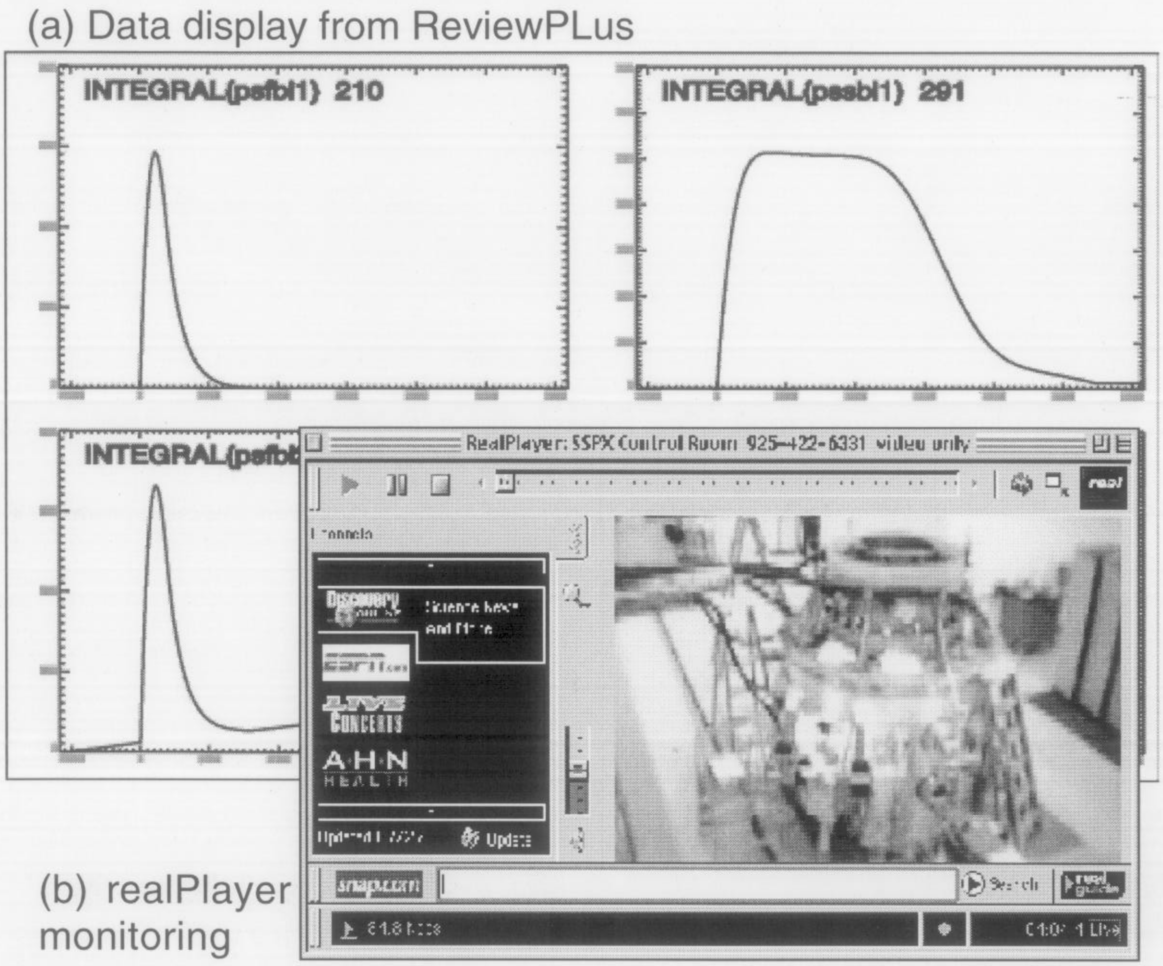

Figure 2 SSPX (a) ReviewPlus analysis and display and (b) RealVideo monitoring.

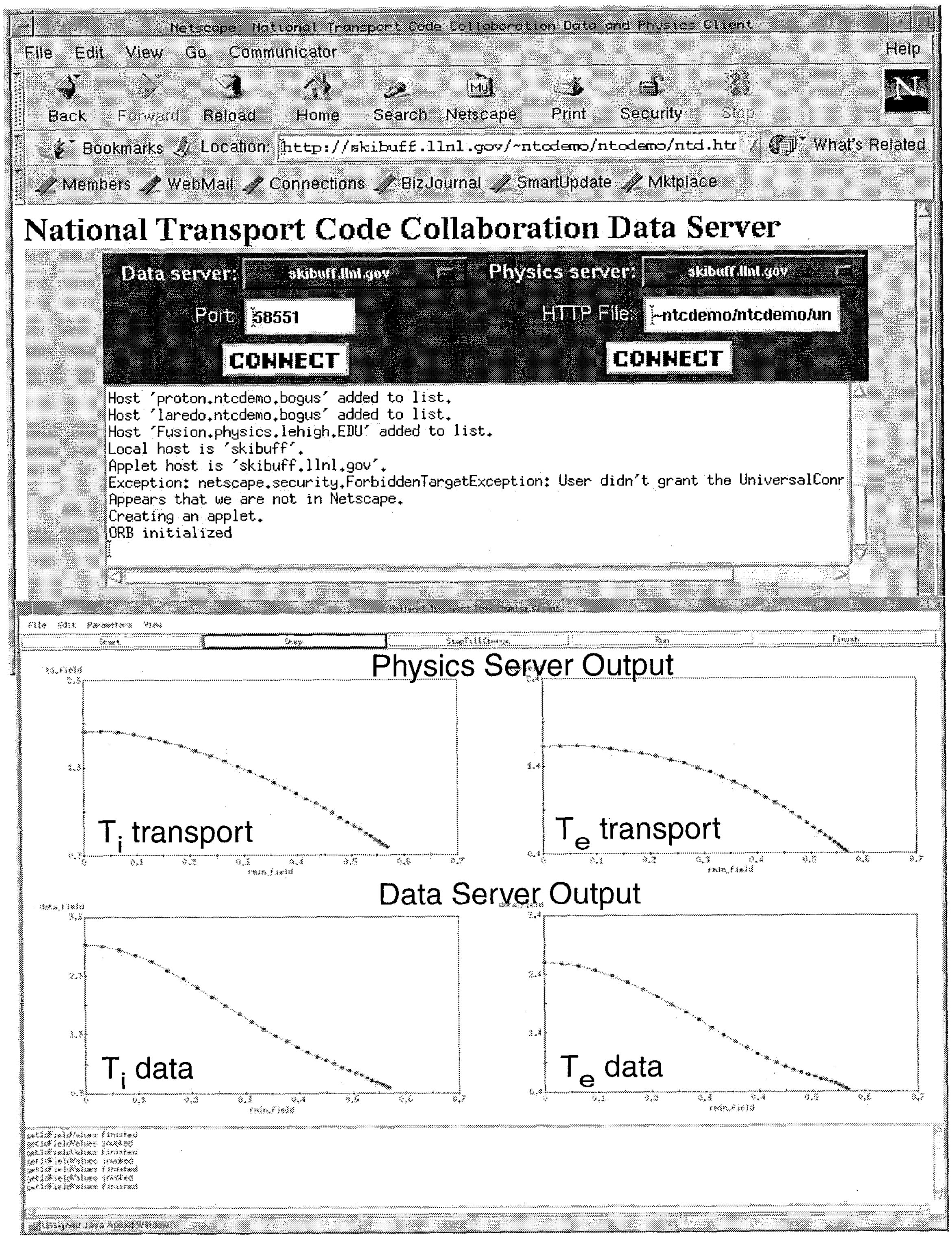

Figure 3. NTCC (a) client application for accessing (b) physics server for transport calculations and data display.

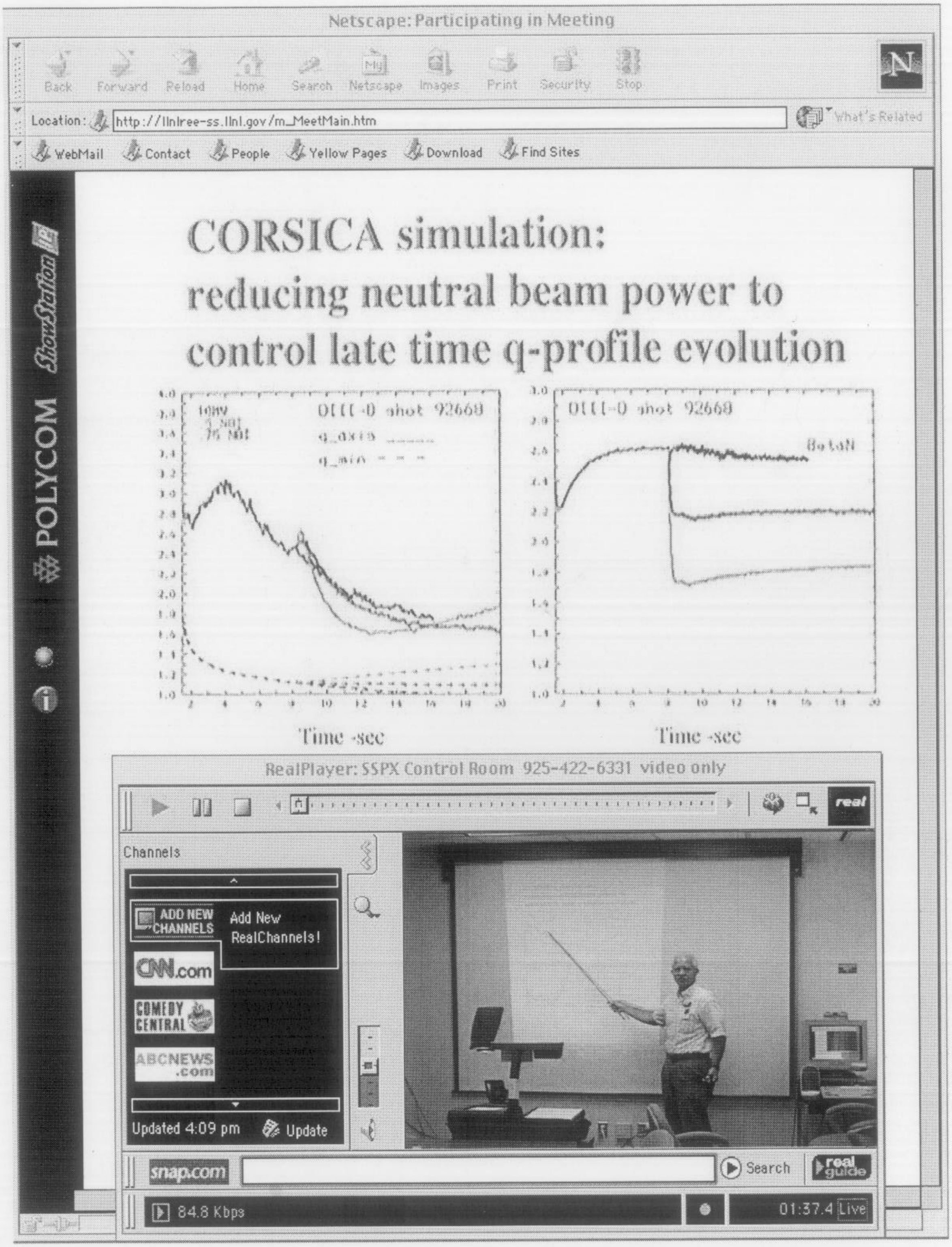

Figure 4 ShowStation and RealVideo technologies for off-site meeting support: (a) presentation output to web browser and (b) RealVideo view of meeting.

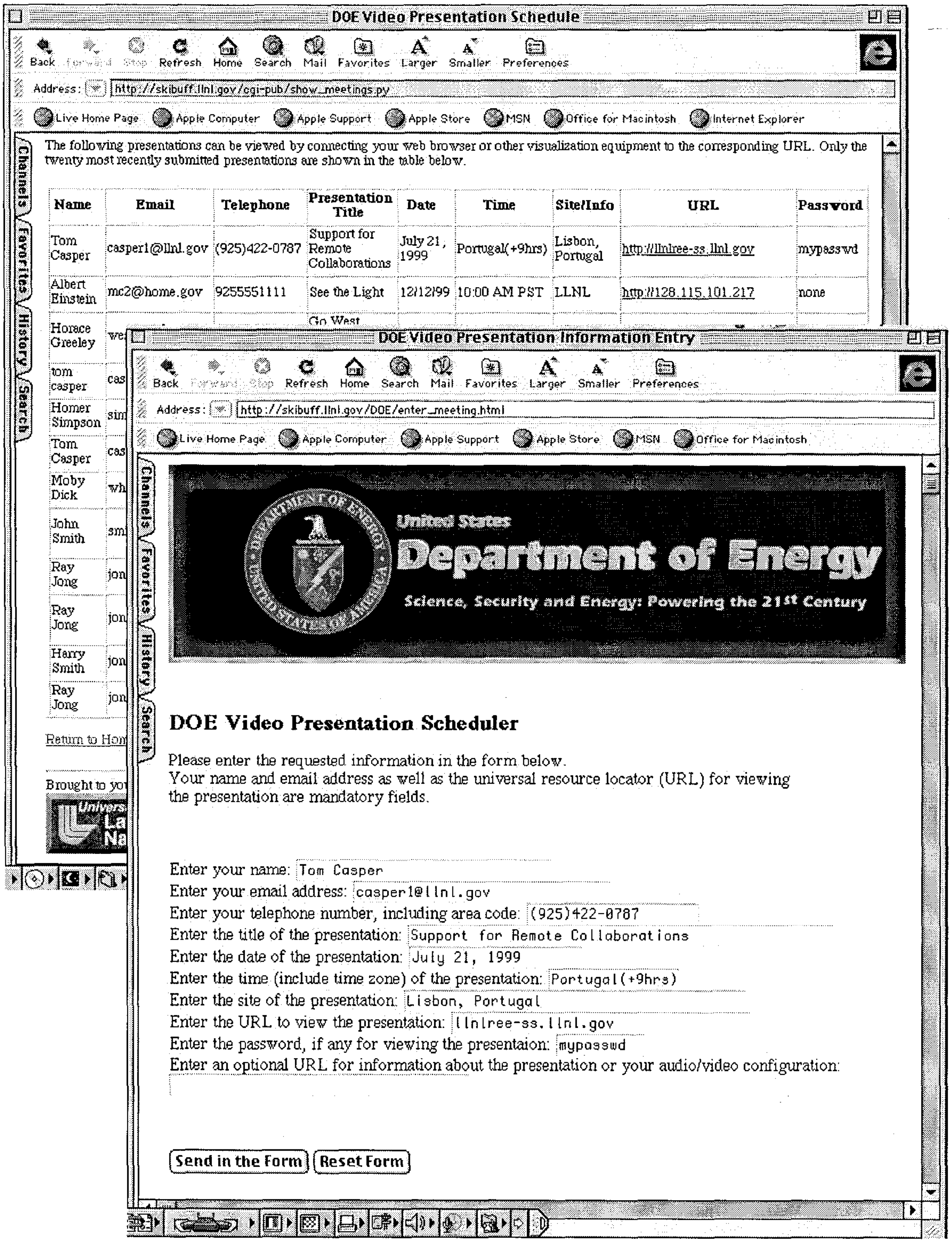

Figure 5 Web browser applications for meeting support: (a) meeting entry and (b) schedule.